

DOCUMENT RESUME

ED 221 841

CS 006 836

AUTHOR Davison, Alice
TITLE A Systematic Definition of "Sentence Topic."
Technical Report No. 264.
INSTITUTION Bolt, Beranek and Newman, Inc., Cambridge, Mass.;
Illinois Univ., Urbana. Center for the Study of
Reading.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
PUB DATE Oct 82
CONTRACT 400-76-0116
NOTE 56p.; Earlier versions of this paper were presented
at the NWAWE meeting at the University of
Pennsylvania (10th, Philadelphia, PA, October 1981)
and as colloquia at the University of Arizona and
Michigan State University.

EDRS PRICE MF01/PC03 Plus Postage.
DESCRIPTORS *English; *Grammar; Reading Comprehension;
*Sentences; *Sentence Structure; Syntax
IDENTIFIERS *Subject (Grammar)

ABSTRACT

The grammatical subject of sentences in English is regularly but not invariably perceived as the sentence topic. Attempts to express this regularity as a rule of grammar are frustrated by the numerous cases in which there is no topic or some other referring expression is the topic. An alternative account is one in which sentence topic is inferred on the basis of properties along two linguistic dimensions. The first is syntactic markedness: the more marked the syntactic structure, the more strongly the subject or other salient noun phrase (NP) is perceived as the topic. The second scale involves the referential explicitness of a possible topic NP. The NP types most compatible with being sentence topics are the ones which identify the most clearly their intended referent. In both cases, the definition of topic is based on the paradigmatic opposition of elements in English, serving as a "context" for all utterances. More marked structures and most explicit NPs have a communicative function of identifying what the sentence is about, relevant to the processing of syntactic-semantic structures in discourse. (Author/HTH)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

ED0221841

This document has been reproduced as
received from the person or organization
originating it.
Minor changes have been made to improve
reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

CENTER FOR THE STUDY OF READING

Technical Report No. 264

A SYSTEMATIC DEFINITION OF SENTENCE TOPIC

Alice Davison

University of Illinois at Urbana-Champaign

October 1982

University of Illinois
at Urbana-Champaign
51 Gerty Drive
Champaign, Illinois 61820

Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, Massachusetts 02238

The research reported herein was supported in part by the National Institute of Education under Contract No. HEW-NIE-C-400-76-0116. Earlier versions of this paper were presented as a paper at the Tenth NWAWE meeting, University of Pennsylvania, October 1981, and as colloquia at the University of Arizona and Michigan State University.

5086836

7

EDITORIAL BOARD

William Nagy and Stephen Wilhite
Co-Editors

Harry Blanchard

Anne Hay

Charlotte Blomeyer

Asghar Iran-Nejad

Nancy Bryant

Margi Laff

Avon Crismore

Terence Turner

Meg Gallagher

Paul Wilson

Abstract

The grammatical subject of sentences in English is regularly but not invariably perceived as the sentence topic. Attempts to express this regularity as a rule of grammar are frustrated by the numerous cases in which there is no topic or some other referring expression is the topic. An alternative account is proposed in which sentence topic is inferred on the basis of properties along two linguistic dimensions. The first is syntactic markedness; the more marked the syntactic structure, the more strongly the subject or other salient NP is perceived as topic. The second scale involves the referential explicitness of a possible topic NP. The NP types most compatible with being sentence topics are the ones which identify the most clearly their intended referent. In both cases the definition of topic is based on the paradigmatic opposition of elements in English, serving as a 'context' for all utterances. More marked structures and most explicit NPs have a communicative function of identifying what the sentence is about, relevant to the processing of syntactic-semantic structures in discourse.

A Systematic Definition of Sentence Topic

In this paper I will propose a characterization of sentence topic for English. Much previous linguistic work has been done to systematize the intuitions of native speakers concerning how they perceive what a sentence is about, insofar as Chicago is north of Champaign can be understood to be about Chicago. Interpretations of this kind often strongly influence judgements of the meaning and well-formedness of sentences, and for this reason the factors contributing to the perception of sentence topic need to be defined and factored out of grammatical description. Even though the notion of sentence topic is often bound up closely with grammatical aspects of a sentence, including word order, morphology, syntactic structure and the semantic content of constituents, it is difficult to define sentence topic using the same formal definitions which describe regularities of syntax and morphology, or truth conditions which define meaning. The characterization proposed in this paper will not be part of the grammar of English, though it will be based on and derived in part from the grammar of English. More specifically, it will be based on the properties of the grammatical system of English, and what is relevant to this definition is the system of syntactic oppositions and semantic contrasts which are defined by the rules of grammar in English. In the discussion which will follow, the syntactic system referred to will be a transformational grammar in a very broad sense, the specific properties of which will not be crucial except that syntactic relations among 'constructions' are statable in some way. For example, the preverbal NP in a 'passive construction' corresponds in grammatical role with the postverbal NP in an active transitive construction.

Given a sentence, especially without a discourse context, it is difficult if not impossible to tell what a sentence is 'about,' and different people may give very different answers. Given two sentences or more with corresponding contents and different syntactic form, or similar syntactic form and different contents, it is possible to distinguish how the contrasting sentences may differ in what they are 'about.' Active and corresponding passive sentences may differ in just this way, with somewhat more consistency of judgement among speakers of the language than for single sentences in isolation. What is revealed by contrasts between otherwise similar sentences will be an important basis for the proposal I will make here. As the contrasts consist of differences of syntactic structure and of semantic content of some constituents of a sentence, I will confine the characterization of what a sentence is 'about' to the notion of what the topic of a sentence is, not what the topic of a discourse is. I will discuss the relation between the form and content of a sentence, and the range of possible sentence topics which these properties determine.

It is important, as Reinhart (1981) shows, to distinguish between discourse and sentence topics. She points out that that while discourse and sentences consisting of sentences, or linguistic units, may have a topic, it is also true of non-linguistic things such as pictures, or films, that they have a topic. Even in a linguistic discourse, the perceived topic need not be a part of the discourse, corresponding to a specific sentence or sentence constituent. In contrast, a sentence is necessarily a linguistic unit, and its (sentence) topic is a constituent of it, such

as a subject NP or dislocated phrase. Of course, there can be some connection between sentence and discourse topics. If the discourse consists of a single sentence, then the topic of the sentence is identical to the topic of discourse. At different points in a richer and more complex discourse, the topic of a sentence may coincide with the topic of a discourse. It is possible for this reason to speak of the 'fit' between sentence topic and the discourse context of the sentence. This relation between sentence and discourse is, however, more clearly perceived when there is a lack of fit, when the preceding discourse clashes with the sentence topic defined by properties of sentence structure and content. There need not be a high degree of matching between context and sentence structure, and I will not assume that context in any way determines the structure of a sentence.

Following Reinhart (1981) I will take the term sentence topic to refer to what the sentence is about, in the sense that a sentence is understood to predicate some property of a given entity. This definition has to be supplemented by a definition of how the constituent referring to sentence topic is isolated from the rest of the sentence. The body of the paper will be concerned with how sentence structure serves to define sentence topics. 'Aboutness' in itself does not constitute a reliable criterion for defining a topic. In fact it is not a satisfactory definition; it is merely preferable to alternatives, such as 'old information.'¹ Some of the failure of previous efforts has been that the discourse function of topics was identified with the criteria for defining what it is. If the function varies from discourse to discourse, then the

definition is too vague, and some definitions--taken by themselves--suffer from this flaw. Others which are more precise fail because of the number of exceptions arising from contextual factors, the effects of more than one variable being involved (including syntactic, semantic and phonological factors) and the fact that sentence topic is not a necessary grammatical component of a sentence, such as subject and verb.

In addition to there being disagreements about what a sentence topic is, there are also many different criteria used to pick out the constituent which is likely to be perceived as sentence topic in a given sentence. Some of the definitions of sentence topic focus on the pragmatic properties of the elements in a sentence: what a given NP refers to and what knowledge about the referent is assumed by the speaker to be shared by the speaker and hearer. This knowledge may be manifest, if there is explicit reference to an individual or state of affairs in the preceding discourse. Such shared knowledge is often called known or old information (Chafe, 1974, 1976; Haviland & Clark, 1974). Prince (1980) refines this notion to include, under the general term evoked, information which is implicitly shared as well as explicitly mentioned. The fact that a linguistic expression refers to something previously mentioned or already known may be marked phonologically, with low pitch on anaphoric pronouns (Chafe, 1976). This phonological mark defines that part of the sentence as not part of the sentence receiving prominence in the normal sentence intonation contour. This is especially the case of the element with low pitch is in the last major constituent of the sentence, where it would normally receive the most phonological prominence. (Multiple constituents with low pitch,

indicating old information may isolate a single constituent as contrastively pitched, Schmerling, 1974).

Other features of sentence topic are morphological and syntactic. The grammatical subject of a sentence in many languages is identified with the sentence topic (Li, 1976). Li and Thompson (1976) note that if topic is not regularly marked in a language by a special particle such as Japanese wa, then subject marking also takes on the discourse function of marking sentence topic. Subject marking involves word order, position defined with respect to the verbal complex, as well as morphology, verbal agreement and nominative marking on the NP. It is interesting to relate Li and Thompson's characterization of the patterns of subject functioning as topic in many languages with Prince's finding that 90% of the grammatical subjects in a discourse referred to things evoked by the prior context, in a sample of spontaneously elicited conversation (Prince, 1980).

Other factors which are associated with sentence topic-hood are the referring properties of a Noun Phrase. These may include specificity of reference, or reference determined independently of any other expression in the sentence (Keegan, 1976). Strawson (1971) has made the well-known distinction between subject and predicate Noun Phrases on the grounds that failure of a NP, such as The present king of France, to refer to anything has different consequences for subject and non-subject NPs. A subject NP which fails to refer renders the sentence it is in without truth-value, in Strawson's view, while the failure of a non-subject NP to pick out a referent in the world of discourse in which the sentence is uttered merely renders the sentence false. If, as I will assume, it is normal in English

for subjects to be sentence topics, then the distinction which Strawson makes could also apply to a difference between sentence topic NPs which fail to refer and other NPs which fail to refer. The consequences of failure would seem to be the same: If the sentence topic defines what the sentence is about, then it follows that it is hard to judge the truth value, or to assign a definitive truth value, to a sentence which turns out to be about nothing. This is especially true of sentences which purport to be about something, and predicate particular properties of the topic.² A sentence which is pragmatically contradictory in this way could not be assigned truth value.

I will take as an operational definition of sentence topic that it is a Noun Phrase constituent of the sentence of which it is the topic, and that the topic represents what the sentence is about; following the arguments in Reinhart (1981). There are some immediate objections to this definition, since there are sentences with no readily identifiable topic, or for which there is disagreement about which of many possibilities could be the sentence topic. I will also propose and discuss in more detail in later sections why sentences do not have multiple topics. I will also use as a basis for my proposal the definition of grammatical subject as sentence topic, following Li and Thompson (1976), though some modifications and extensions of this definition will be necessary. There are several immediate objections to this characterization also: (a) there are sentences with grammatical subjects which either lack a perceptible topic, or if they have one, it is not the subject; and (b) as in the first set of objections above, if there are several possible topics, how can topic function be

uniquely assigned to the grammatical subject? Since there is only one surface grammatical subject per clause, then there can be at most one sentence topic per clause, perhaps none, in the case of sentences with imperative form which lack overt subjects. Finally it may be objected that only one NP is assigned topic function, yet a sentence may have many clauses, each of which may contain a grammatical subject.

A Proposal for the Definition of Sentence Topic

The answers to these objections will emerge from the discussion of cases to follow. Before dealing with them specifically, I will illustrate how a definition of subject as topic would fit an ideal case, in (1); this definition will be given in more detail in (7).

- (1) a. (Context) The investigation continues.
 b. On Tuesday, the police arrested two suspects.
 (Background) (topic) (focus)

The grammatical subject of (1)b, the police, may be perceived as the topic of the sentence. It may be linked by real-world knowledge and semantic connection to the investigation in the context sentence (1)a, and so the topic of (1)b may be characterized as evoked in Prince's (1980) term.

The object NP, two suspects, receives the intonation peak of the declarative sentence contour in (1)b, and thus represents the focus of the assertion, cf. Chomsky, 1971, and Chafe, 1974. Hence it may be perceived as new information introduced into the discourse. A sentence which followed (1)b might contain a sentence topic linked in some way to the focus constituents of (1)b. The preposed adverbial on Tuesday has another status, neither topic nor focus. Assuming that the 'normal' or 'unmarked' position of adverbials in a clause is after the subcategorized elements of the VP,

then in sentences like (1)b a preposed adverbial occurs not only outside of its normal position, but also outside of the position where it would receive intonational prominence. Such adverbials in preposed position are generally perceived as background or assumed information.

Sentence topic has been identified by some writers with old or given or assumed information (Haviland & Clark, 1974; Chafe, 1974). But the case of preposed adverbials shows that for a constituent to convey old or background information is not sufficient to cause the constituent to have the role of topic in a sentence. Topic-hood also requires, as we will see, a certain kind of prominence or perceptual salience, which is not the same thing as sentence focus or intonational prominence (cf. Olson & Filby, 1972), and is also not conveyed by a preposed adverbial phrase. Even if preposed adverbials are given prior mention in a context sentence, so that the entity referred to in the adverbial phrase of a succeeding sentence could be perceived as a sentence topic, the sentence form itself prohibits this. A NP is not readily linked with a discourse antecedent if it is initial in the sentence but is not the grammatical subject (Davison & Lutz, to appear).³ The definition of theme or topic as the initial element of a sentence in Prague School work such as Firbas (1966), and in Halliday (1967) is too broad, since it includes not only preverbal subjects, as it should, but also preposed adverbials, which are not possible sentence topics. Instead, for a constituent to be perceived as sentence topic requires both preverbal position and grammatical subject marking--though these properties will be subject to revision and modification in the discussion to follow.

In this paper, I will identify grammatical subject properties with the potential for being perceived as sentence topic. Here I want to examine how the notion of sentence topic can be a pragmatic one, not part of the grammatical system of a language or part of the syntactic structure of a sentence, and still be associated with grammatical systems and syntactic constituents. Since the perception of a NP as sentence topic is so subjective and variable, there must be at least several factors which influence perception. In the next section, I will consider a range of cases, in which sentence properties are systematically varied, as are semantic properties of NPs in the sentences. Since the notion of sentence topic most clearly emerges when two similar sentences are compared in isolation, or when there is a lack of fit between a context sentence and a succeeding sentence, both the discourse context and the choice of preferred continuation will be important factors.

Marked and Unmarked Cases

Let us suppose that speakers of English regularly but not invariably make the inference that grammatical subjects are sentence topics, all other things being equal. This inference could be a case of generalized (i.e., not specific to one context or occasion) conversational implicature (Grice, 1975). This inference could be based on Grice's Maxim of Manner, which urges the speaker to use the most appropriate form for the message uttered; if a particular NP is placed in subject role, out of all the possible grammatical roles which the grammar allows, then an inference may be drawn about the speaker's intention to place the NP in a grammatically prominent position. If subject position is indeed a salient one, by virtue of its

grammatical relation to the verb and potential of nominative marking, then Grice's Maxim of Relevance might also be invoked; grammatical prominence is an indication of discourse prominence, and the speaker is indicating how the sentence containing a subject NP is to be related to what went before or is known. If the identification of subject with topic is not a case of Gricean inference, then it could be described as a discourse-processing strategy of some sort, equally independent of the grammatical system. In the interest of not multiplying descriptive entities, I have taken the choice of including topic perception as among the inferences based on Grice's Maxims, which are operative in any case. The exact account of the link between subject and topic is not crucial, so long as it is stipulated that the association is not a grammatical one, not an invariant and arbitrary association of the kind which is true of linguistic signs in general, such as the association of /-s/ with plurality in English.

If the association of subject properties and perception as sentence topic is both pragmatic and regular, it is cancellable under some circumstances. For example, the sentence topic of (2)b should be a kid:

- (2) a. (context) An awful thing happened to my brother.
 b. A kid stole Jim's wallet.
 c. He was furious.

Yet I think most people would agree that a more likely topic is Jim, even though it is a possessive NP within the object NP. But even though Jim is not a subject NP, it refers to an individual evoked by the context (2)a, by the phrase my brother. Continuity of reference is possible in the sequence my brother . . . Jim, but not in the sequence my brother . . . a kid, because the indefinite article indicates that there was no previous

mention. The occurrence of a first-mention NP in subject position and the identification of a non-subject with the topic in this example shows that the Gricean inference equating subject and topic may be cancelled without serious contradiction, or interruption of the discourse. There is nothing strikingly odd about the discourse in (2), with (2)c a continuation of reference to my brother.

Next let us turn to a more 'marked' syntactic structure, as in the passive sentence of (3)b. Syntactic markedness will be discussed later in more general terms, but here the passive structure may be regarded as more marked by comparison with the equally available⁴ active structure. The passive form reflects less directly than the active form the underlying grammatical relations of object, agent, experiencer, etc. The use of surface structure information, category and order to assign logical structure requires more steps, and specific rather than general strategies (Bever, 1970) in the case of the passive structure. In (3)b, the grammatical subject is Jim, but the context (3)a evokes a kid:

- (3) a. School must be out now.
- b. Jim_i was robbed by a kid_j on Friday.
- c. He_j ran away.
- d. He_i was furious.

Compare the same sentence, (3)b, with a different context sentence in (4):

- (4) a. An awful thing happened to my brother_i.
- b. Jim_i was robbed by a kid_j on Friday.
- c. He_j ran away.
- d. He_i was furious.

The sequence of sentences in (3) might seem somewhat less connected than the sequences in (2) and (4)a, b. In (4), the expression my brother is

in focus position in (4)a and is linked referentially to the NP Jim in subject/topic position in (4)b. There is no interruption of expectations about topic in the transition from (4)a to (4)b, as there is in the transition from (3)a to (3)b:

One test for the perception of topic in these last two cases might be the preferred reading for he in (3) and (4)c and d. If these sentences are possible continuations of the discourse, and if he is likely to be identified with a salient element in previous discourse, such as a topic (cf. Kantor, 1977), then the preferred referent of he would be more definitely determined the more strongly the previous sentence (3)b or (4)b determines a topic, with prior discourse information as well as structure and NP content.

My intuitions are that the choice is indifferent in (3) and that (4)d is preferred over (4)c. If so, then the inference that subject equals topic is cancelled in (3) and maintained in (4). Neither choice produces an illformed sequence; The difference is just that the sequence (4)a, b is more 'connected' than the sequence (3)a, b. In other words, there is a better 'fit' between (4)b and its context (4)a than between (3)b and its context (3)a.

Next, let us look at a still more 'marked' sentence form, a passive sentence whose grammatical surface subject is not a 'real' object, or patient, but rather the object of a preposition not subcategorized by the verb. If the prior context suggests that the subject of the sentence (6)b is the topic, by introducing a possible antecedent, for example, then the sequence seems well-formed. But if the context suggests that

the topic is the referent of some other expression in the sentence, then the sequence appears less well-formed. For example, the context (5)a suggests that the topic of the next sentence will be the agent phrase, rather than an expression of location. But the locative object is in subject/topic position in (5):

- (5) a. Sam sat over there on the right.
- b. ?? This chair was sat on by Fred.
- (6) a. You may wonder what happened to the furniture.
- b. This chair was sat on by Fred.

The match between context and inferred sentence topic in (6) is a relatively good one, while the match in (5) between the same sentence b and a different context is much less good. In the sequences given in (2) - (4), we have seen that similar manipulations of the contexts of active and passive sentences of the usual kind make much less difference than in the contrasting cases illustrated in (5) and (6). Here it appears that the inference that subject = topic applied to (5)b is not easily cancellable. The sentence structure of (5) - (6)b can be characterized on independent grounds as more 'marked' than active and ordinary passive structures because the surface subject is neither an underlying agent nor patient in semantic structure, since sit is not used here as a transitive active verb.⁵

The general point I want to make here is that the syntactic structure contributes to the definition of what is perceived as sentence topic, but some variability is to be expected. The amount of variability is restricted in some cases, however. We may summarize the results so far for defining how grammatical subjects are equated with sentence topics:

(7) The subject-topic relation

- a. If it is a normal conversational inference that the surface subject of a finite main clause is the topic of the sentence, then the inference is cancellable in unmarked sentence structure, and less easily cancelled in marked syntactic structures.
- b. Syntactic markedness is determined by how directly the surface syntactic relations reflect underlying semantic relations of grammatical role or function/argument structure. Markedness might be measured in how well simple equations of surface and underlying structures work (cf. Bever, 1970; and Wanner & Maratsos, 1978, on surface relations).

Markedness of syntactic structure can also be characterized in terms of the class of syntactic structures which are defined as expressing equivalent syntactic relations by the rules of syntax and semantic interpretation. There are structures which bear close similarity to one another in the grammatical relations expressed, such as actives and passives of both the usual and unusual kind, but out of these, one reflects underlying relations quite directly--the active structure--and one is a very indirect expression of grammatical relations--namely, the 'peculiar' passive of (5)b.

Properties of Referring Expressions and Topic Function

In the preceding cases, and others to follow in later sections, we see that sentence structure can be the basis of inferences defining the sentence topic. In this section, it will be shown that there is another important factor operative in the perception of sentence topic function for a referring expression. This factor interacts with the grammatical features just discussed, and may reinforce them, but it also may be

independent. What is involved is the information semantically expressed in the NP which defines what the NP refers to, and how clearly this information defines the referent. For example, in the sentence (2)b, the grammatical subject is a kid, but it is not a good possible topic. This NP is indefinite, conveying that the identity of the referent is not known to the hearers, probably not to the speaker either. Hence it has less clear reference to an individual in a discourse context than the definite NP my brother and the proper name Jim. A NP will therefore be likely to be perceived as sentence topic if it is in grammatical subject position or if it contains information defining its referent clearly. In example (2), we saw that the inference based on grammatical information could be cancelled by a combination of discourse information and referring properties of NPs in unmarked structures such as active sentences, but this will not be the case with more marked structures, such as sentences in which Passive or raising rules apply.

An extreme example of how referential properties of NPs interact with topic function is the class of sentences containing the expletive there, as in there are 3 points to consider, 'Existential' there refers to no entity in the discourse context, yet it occurs in preverbal position; and it counts as a grammatical subject for the purposes of the rules of Auxiliary Inversion and Raising rules. It lacks only the subject property of verbal agreement in finite clauses. As a subject of a sort, there may be assigned the role of sentence topic, so that there is what the sentence is about. But the NP there, which is not even a locative deictic in this case, gives no information leading to the identification of a referent.⁶

Either the inference is maintained, so that topic is vacuous and the sentence is not about a particular referent, or the inference is cancelled, allowing some other NP with stronger referential properties to be perceived as topic. But the remaining possibilities in the sentence are either indefinite and non-topics, according to Milsark's (1977) characterization of 'existential' there, or they are proper names and definite NPs in focus position, in the 'list' interpretation (cf. Rando & Napoli, 1978). NPs in focus position are generally not perceived as topics because they are interpreted as new information, part of what is predicted about a topic.

Some of these features of there are illustrated in the contrasts shown in (8) and (9):

- (8) a. We have to keep the booze hidden from Uncle Harry.
- b. Ten cases of beer are stacked in the garage.
- c. ? There are ten cases of beer stacked in the garage.
- (9) a. I had to park my car in the driveway.
- b. ?? Ten cases of beer were stacked in the garage.
- c. There were ten cases of beer stacked in the garage.

In the first example, (8), the version of the sentence without there is preferred over the one with there, (8)c. The indefinite NP ten cases of beer is a sub-instance of previously mentioned booze, which thereby constitutes a discourse antecedent for ten cases of beer. The version of the sentence, (8)b, is preferable because it places this NP in subject position where it may be perceived as a topic without any contradiction from context. The reverse is true in the context (9)a, which does not introduce a discourse referent for ten cases of beer. Hence the version of the following sentence which is preferable is (9)c, rather than (9)b. The presence of

there in (9)c keeps ten cases of beer out of subject position, so that it is not interpreted as a topic. The version (9)b where it does appear in subject/topic position implies conversationally that the existence of the beer is already known, and further inference is necessary to connect (9)b with the context (9)a (cf. experimental results in Haviland & Clark, 1974, for definite NPs introduced without a discourse referent). One important consequence of the fact that there has null reference is that as an overt constituent it keeps certain NPs out of subject position, where they would otherwise be perceived as topics, and so interfere in the perception of the appropriate connections in discourse.

Some NPs which should not be perceived as topics are those which lack a discourse antecedent, and are not intended to have one. Such NPs are often indefinite NPs in the non-specific interpretation. We may contrast such banal instances of there sentences as (10)a and (11)a with the same sentences without there.

(10) a. There is a fly in my soup.

b. ? A fly is in my soup. --Which one?

(11) a. There doesn't seem to be a waiter here.

b. ?? A waiter doesn't seem to be here, We want to order lunch.

The lack of there in (10)b and (11)b allows the unwanted inference that the indefinite NP indicates a sentence topic, and that the speaker has a specific instance in mind.⁷ An indefinite NP in subject position tends to receive the specific interpretation over the non-specific interpretation by virtue of the topic inference. If the NP represents a topic, it indicates what the sentence is supposed by the speaker to be about. If the speaker has no particular thing in mind, then the sentence is not a

useful addition to discourse and the sum of knowledge in that situation (except of course where it is understood that detailed knowledge is impossible, as in discussions of the origin of the universe, etc.). By Gricean assumptions, the speaker is purporting to make a useful contribution, so that the sentence is given the strongest interpretation consistent with everything else in the situation.

The same effect of displacing a NP from subject position also has the consequence of placing a definite NP in final or focus position, where they are perceived as 'asserted' rather than stipulated.

- (12) a. What shall we use for firewood?
 b. There's the chair with the broken back.
 c. There was a chair with a broken back stacked in the garage.
 d. ? There was the chair with the broken back stacked in the garage.

The sentences (12)b-d are meant to be possible continuations from (12)a, which does not introduce a direct discourse referent for a/the chair with a broken back. The continuation in (12)b is possible, with the definite NP in the position where it receives the intonation peak. Similarly, the indefinite NP in (12)c is possible, although it is not in the focus position; that position is occupied by the locative expression. But both of these are clearly to be preferred over (12)d, which has a definite NP, a possible topic by virtue of its referential properties and also by virtue of its not being in focus position. Hence (12)d is odd in the context established by (12)a, though perhaps not in other contexts.

The referring properties may be arranged in the following hierarchy, according to how distinct they are in reference and how easy it is to pick out the intended referent according to the information give (or inferrable) in the semantic contents of the NP.

- (13) a. (i) Proper names: Harry, Miss Madeira, France
 (ii) Referential uses of definite descriptions (vs. attributive uses b. (iii); cf. Oonellan, 1971):
Smith's murderer, referring to a certain individual.
 (iii) Referentially 'transparent' descriptions: Oedipus did not recognize his mother.
 (iv) Specific indefinites: a (certain) policeman.
- b. (i) Generic NPs: The owl, a typical small town, students.
 (ii) Abstract NPs of time and place: from 1492 to 1588, 5 miles.
 (iii) Attributive definite descriptions: Smith's murderer (whoever it may be).
 (iv) Referentially opaque descriptions: Oedipus wanted to marry his mother.
 (v) Non-specific indefinite NPs: a policeman (no particular one).
- c. (i) Non-referring NPs in idiom chunks: tabs, the jig.
 (ii) superlatives: the slightest noise.
 (iii) Generic NPs: any doctor.

The hierarchy exemplified in (13)a, b and c goes from the best cases a to the worst cases c, with no particular internal relations in each group. The proposal is that NPs of the type illustrated in (13)a make excellent topics, and those in (13)c make mediocre to poor topics, while those in (13)b fall between the two extremes.

This hierarchy actually comes from a study of Raising to Object (Borkin, 1974), and the gradations were established independently of the notion of sentence topic. Borkin recorded the preferences of English speakers for raised or finite, unraised complement subjects, varying the NPs raised, the form of the complement (infinitive, as or \emptyset) and predicate type. The hierarchy in (13) shows what type of NP best undergoes Raising

to Object, namely those in (13)a, and what NP types generally produce strange or less preferred raised structures. Some specific examples of these interactions of NP type and sentence structure will be given in later sections. But it should be noted here that the NP types which are highest in Borkin's hierarchy are also NPs with the clearest or most easily determined referents. They are the ones about which the speaker, and perhaps also the hearer, have the most information, perhaps derived from first-hand knowledge. The intermediate cases involve entities not directly known to the speaker, known only by properties; they may also be abstract entities which cannot be individuated or directly perceived. The worst cases are ones which require exhaustive search; for superlatives, the unique referent is the noise which is the slightest one, etc. Any generics have no specific referent, but also do not allow consideration of all cases collectively, or cases with reference to a set, since any is different from all, every, each and is always a non-specific indefinite (cf. Vendler, 1967, Davison, 1980a). Nonreferring expressions which are part of idioms are misleadingly like referring expressions, having no reference though the whole has a meaning derived non-compositionally.

This hierarchy interacts with the conversational inference defining subjects as topic, introduced earlier. If a NP high up in the hierarchy (13) occurs in subject position, it will combine with the structurally based inference, and will be a clearly perceived topic. An example of this combination is illustrated in (1)b, where the grammatical subject is the definite NP the police: (1)b On Tuesday, the police arrested two suspects. In such cases, the referential properties of a NP which make its referent clear or easy to identify reinforce the inference that the subject NP is topic. Conversely, if the subject NP belongs to a type at

the bottom end of the hierarchy, then the structurally based inference about subjects may be cancelled, especially if there is another NP in the sentence with better qualities as a possible topic. The sequence in (2) and (14) are examples of this combination:

(14) Anything annoys Charlie.

The topic of (14) is more likely to be perceived as Charlie than the grammatical subject anything. The use of a proper name like Charlie implies conversationally that the speaker has a particular individual in mind, assumed to be known to the hearer as well. If the sentence is about Charlie, it is a clearer contribution to the discourse than if it is about anything (at all).

The referring properties of NPs in (13) may serve to cancel the inference equating grammatical subjects with topic, just as context may. But what if the inference is not cancellable? We have seen that it is less easy to cancel the inference in some structures which are more marked than others, such as the prepositional passive sentences discussed in Davison (1980b). A hierarchy of grammatical roles was proposed there, representing preferences for NPs which could be promoted to subject role. NPs referring to abstract or unindividuated things generally make poor subjects of passive sentences; these include expressions of time (except for events: this experience, etc.), destination, cause and manner.

Definite NPs referring to concrete, individuated entities make generally acceptable subject of passive sentences formed from prepositional objects. This hierarchy is in effect part of the hierarchy in (13), the main difference just being that there are different restrictions on what can occur as a clause subject as opposed to a prepositional object. Further,

there are differences in how marked different constructions are, and thus the degree to which they tolerate 'poor topic' NPs may differ.

Degrees of Syntactic Markedness

But what if there are structures which are so marked that the conversational inference of topic function is not cancellable, making the inference conventional (Grice, 1975)? In such cases, the marked structures and the NPs low in the hierarchy of (13) should be incompatible. This is the case for Raising to Object sentences and other constructions which will be discussed in this section, taking incompatibility to refer to general tendencies rather than absolute violations of grammaticality. Part of the justification for making topic definition distinct from the grammatical system of a language is that topic definition, even in highly marked constructions, is still somewhat variable. As Borkin (1974) shows, it is possible in some cases to combine raised structure with a 'bad' NP type, and produce a relatively acceptable sentence, provided that there are compensating factors such as a comment-like predicate. Conversely, it is possible to combine a 'good' NP with other less favorable factors and produce a strange or deviant sentence. The structure associated with Raising to Object can be characterized as marked in that the surface structure relations do not match the underlying ones, and are more misleading than helpful in ordinary parsing routines; cf. (25)b discussed below. The subject of the complement clause has object marking, suggesting that it should be an argument of the preceding verb of the main clause, rather than the verb of the complement. The clause boundary of the complement is obscured, by comparison with the finite structure. The

higher clause is pragmatically 'transparent,' because of the meaning of the verbs which govern Raising to Object. These all have to do either with belief and knowledge, or with assertion and report, and so are often redundant in an illocutionary act of asserting or questioning: declare, regard, suppose, report, etc. These factors combine to define the subject of the complement clause as the topic of the sentence as a whole, and there is experimental evidence that this is the case in Davison and Lutz (to appear).⁸

The effects of the raised structure may be seen if we take a NP type which is mid-way on the scale of (13), an indefinite NP which may or may not have a specific interpretation. If this NP has a discourse antecedent, perhaps a very general one, then the NP has the specific or existential interpretation, in both the raised and unraised form (15)b and c, with (15)a as the discourse context:

- (15) a. There are reports of UFOs in the midwest.
- b. People believe that a flying saucer is spying on Chicago these days.
- c. People believe a flying saucer to be spying on Chicago these days.

Here the mention of a general class, UFOs, establishes a possible discourse antecedent for a flying saucer, an exemplar of the general class. In this case, the context and the sentence structure of (15)c are both compatible with the interpretation of the subject of the complement clause as the sentence topic. The opposite is true in (16):

- (16) a. I don't want to be out on Hallowe'en night.
- b. People believe that a flying saucer is spying on Chicago these days.
- c. ?? People believe a flying saucer to be spying on Chicago these days.

The context in (16)a fails to establish a discourse antecedent for a flying saucer. The indefinite NP in (16)b receives a non-specific interpretation, and the finite complement clause structure is compatible with that interpretation. But the same NP in (16)c must receive a specific indefinite interpretation, because of the marked surface structure which particularly affects the perception of the complement subject, and defines it as topic. The combination of NP type and syntactic structure in (16)c is odd by comparison with the same sentence (15)b in a different form, relative to the context (16)a. The difference of surface structure must have pragmatic consequences which affect acceptability. The only difference between (15) and (16) is the context sentence (a), so that the difference of acceptability of (15)c and (16)c must be due to pragmatic discourse properties combined with surface syntactic structure.⁹

In these examples, the raised NP came from the middle to top of the hierarchy in (13). Let us see how the bad cases of (13) combine with syntactic structures. We have seen that the inference that subjects = topic can be cancelled in unmarked syntactic structures, but not in marked structures. Markedness of structure is not something which can be cancelled in the normal way,¹⁰ so referring properties of NPs cannot override the marked character of syntactic structure in the same way that they serve to cancel inferences based on unmarked syntactic structure, as in the combinations illustrated in (2) and (3). The result of the combination of marked syntactic structure, as in Raising to Object sentences, and referring expressions which make poor topics, is incompatibility, as in the examples below:

- (17) a. They reported that any doctor was able to supply the medicine.
 b. ?? They reported any doctor to be able to supply the medicine.
- (18) a. We believe that the slightest discrepancy is irritating to him.
 b. ?? We believe the slightest discrepancy to be irritating to him.
- (19) a. The records confirmed that tabs had been kept on all the suspects.
 b. ??? The records confirmed tabs to have been kept on all the suspects.
- (20) a. Rocco thought that the jig was up.
 b. ?? Rocco thought the jig to be up.

Here, the raised versions (17)b - (20)b are odd, especially in contrast with (17)a - (20)a, where the raised NP comes from class c of (13). These are NPs without a referent (tabs, the jig) which are part of an idiom which does have a denotation as a whole predicate. Independent null elements such as there as in (8) - (12) and it in weather expressions also undergo raising, but without conflict with syntactic structure. As noted earlier, there moves a non-topic or poor topic out of subject position, as in (10) and (11), while weather predicates lack subject expressions entirely. It is therefore not surprising that these empty NPs can raise, while tabs, etc. cannot. The other NP types which conflict with raised structure are those like generics and superlatives which have a referent, but the unique entity referred to is hard to establish from the semantic information given in the linguistic expression.¹¹

Note that tabs may readily be made the subject of a passive sentence, where it might be perceived as a topic (though it is not likely to be a

topic).. We may then wonder if there are gradations of syntactic markedness, such that raised structures in (15) - (20) are more marked than passive structures, and less marked than others. Raising to Object in particular seems to produce a surface syntactic configuration which is more marked than Raising to Subject, but less marked than NP dislocation structures.

Raising to Object, as illustrated above, prohibits certain types of NP in the complement as subjects, but some NPs of these same types are permitted in Raising to Subject sentences:

- (21) a. ?? We believe the slightest discrepancy to be irritating to him.
- b. The slightest discrepancy seems to be irritating to him.
- (22) a. ??? Rocco thought the jig to be up.
- b. The jig appeared to be up.
- (23) a. ?? The records confirmed tabs to have been kept on all the suspects.
- b. Tabs happened/turned out to have been kept on all the suspects.

Like Raising to Object sentences, sentences in which Raising to Subject has applied also require certain properties in the raised NP. In particular, a discourse referent is required for the NP in subject position if it is a fully referential NP. Note that contrast between (24)b, and (24)c as continuations from (24)a:

- (24) a. Typhoid and cholera were once widespread in the New World,
- b. It seems that many American Indians died in epidemics.
- c. ? Many American Indians seem to have died in epidemics.

The context sentence (24)a defines no specific prior antecedent for many American Indians. The raised structure in (24)c places this NP in subject/topic position, and thus creates a sequence with (24)a as context in which

We may place Raising to Subject structures midway between passive sentences and those in which Raising to Object has applied. A schematic comparison of the three structures, showing their surface structures and the underlying relations which are not directly conveyed, is given in (25), along with Equi-NP Deletion:

- The passive structure in (25)a has the object in preverbal rather than post-verbal position, with a corresponding difference of case marking. The agent is a prepositional object rather than subject. In the Raising to Subject structure, the surface subject of the superordinate clause is the logical subject of the subordinate clause, which is otherwise without a subject. In the Equi-NP structure, the subordinate clause is also null, by virtue of coreference with the superordinate clause subject rather than

via NP movement. In Raising to Object sentences, the subordinate clause subject is marked with accusative case as though it were the object of the superordinate verb, and the clause boundary is obscured, though Borkin (1974) provides evidence that it is not entirely deleted. But it is unclear whether the clause boundary is before or after the accusatively marked NP; cf. Borkin (1974) vs. Chomsky (1981).

Both Raising to Object (25)d, and Raising to Subject (25)b, create surface structures which have ambiguous surface grammatical roles, ambiguous in the way that they indicate underlying subject-verb relations. But the two constructions differ in how 'marked' they are, and this difference is reflected in the facts illustrated in (21) - (23). NP types from the lower end of the hierarchy in (13) may occur more readily in the higher clause subject position in Raising to Subject sentences than in lower clause subject position in Raising to Object sentences. These positions are comparable, because they are the ones whose properties are defined by the syntactic rules in question. It may be possible to explain why Raising to Subject structures are less marked than Raising to Object, by noting the parallelism between (25)b and (25)c, the result of Equi-NP Deletion. In both structures, the subordinate clause lacks an overt subject, and the higher clause subject is identical in reference to the 'missing' lower clause subject, with coindexing brought about in two different ways, movement versus deletion (or movement versus control). Thus while ambiguity of surface structure may contribute to markedness, the absolute markedness values are determined also with reference to other constructions which resemble the marked structure. No such parallel construction is found in English for Raising to Object.¹¹

The Bases of Syntactic Markedness

There are two general principles operating behind the scale of syntactic markedness which compares construction in (25). The first is the Praguean principle of paradigmatic opposition. The rules of English syntax define a class of well-formed syntactic structures. Some of these are related by the fact that the 'same' basic clause structures or combination of elements can be expressed in different configurations of surface elements. These different but related structures may differ in how directly they encode underlying grammatical relations, assuming some continuous mapping from underlying to surface relations. Thus the second principle has to do with the recoverability of semantic relations encoded in grammatico-syntactic terms. Here I have used a way of distinguishing 'direct' versus 'indirect' recoverability by referring to the processing strategies of Bever (1970), but any deterministic statement of a strategy for getting the 'right' answer for the least effort will express the principle I have in mind.

Given these two principles, the consequences for conversationally inferring that subjects are topics are the following: The inference is less easily cancellable, the more marked the syntactic structure is. The consequence for communication, which involves intentional acts of speaking would seem to be the following: Marked syntactic structures are chosen over the equally available unmarked structures for the purpose of conveying additional, pragmatic information.

The application of the term marked to syntactic structures is not new. Its use is implicit or explicit in the work of the Prague School including

Firbas (1966), and in Halliday (1967). The impressionistic use of the term there certainly covers some of the cases discussed here. The difference between these works and the proposal made in this paper lies in how markedness is defined in terms other than observations of what is perceived as topic, and how markedness is linked systematically to other properties of grammar. In a sense, markedness as a way of defining topics mediates between the arbitrary and invariant properties of grammar per se and the variability of perception, as a function of context, and of intentions of the speaker in communication. By relating markedness to Gricean maxims of cooperative conversation (1975), it is possible to explain why in some cases subjects are topics, and in others they are not. Syntactic structure and markedness properties combine with the referring properties of NPs, so that structure and NP contents may reinforce topic-hood, they may conflict, or NP properties may cancel the inference of topic-hood. But syntactic markedness prevents NP properties or contextual information from cancelling the topic inference. Since markedness is defined in terms of the range of structures which form the set of well-formed sentences of the language, the language itself constitutes a kind of context, one which is associated with any utterance of English. The context of a particular utterance may not be sufficient to override this general sort of context.

The Most 'Marked' Syntactic Structures

The most marked topic-defining structures are, not surprisingly, the ones commonly associated with topics. These constructions have been exhaustively analyzed by Gundel (1977). As she notes, certain kinds of

NPs are prohibited generally from occurring in left or right NP phrases, with or without a resumptive pronoun. Predictably, these are NPs of the low end of (13), without reference or without specifiable reference.

Topic phrases are just as strict and sometimes more strict in prohibiting NPs incompatible with Raising to Object:

- (26) a. They supposed that any doctor would know the answer.
 b. ?? They supposed any doctor to know the answer.
 c. ? Any doctor, they supposed (she/he'd) know the answer.
 d. ?? Any doctor, they supposed (him/her) to know the answer.
 e. ?? They supposed him/her to know the answer, any doctor.
- (27) a. ?? We believe the slightest discrepancy to be irritating to him.
 b. ?? The slightest discrepancy, we believe (it) to be irritating to him.
 c. ?? We believe it to be irritating to him, the slightest discrepancy.
- (28) a. We believe nobody to be responsible for the accident.
 b. ??? Nobody, we believe (him/her/them) to be responsible for the accident.
 c. ??? We believe them to be responsible for the accident, nobody.
 d. We believe there to be a problem in administration.
 e. *There, we believe (it) to be a problem in administration.

As the examples in (28) show, dislocated topic phrases place a stronger restriction on their content than Raising to Object structures place on the complement subject. It is possible to raise a NP with null reference, as in (28)a, so long as its lack of referent within a set of possible referents is clear, including negatives and explicitly null elements such as it and there. But topic phrases in structures like (28)b and c cannot be null.

Topic phrases also define sentence topics explicitly, and so can conflict with other features of sentence structure which indirectly define some other NP (and its referent) as a topic. If the NP which is raised to object is also dislocated or topicalized, there is no infelicity or contradiction, as (29)a shows. But if dislocation or topicalization defines one sentence topic and a marked syntactic structure defines another, the sentence is ill-formed or strange pragmatically:

- (29) a. That guy, we believe him/Ø to be the mastermind behind these crimes.
 b. ?? These crimes, we believe that guy to be the mastermind behind them.
- (30) a. We believe him to be the mastermind behind those crimes, that guy.
 b. ?? We believe that guy to be the mastermind behind them, those crimes.

Likewise, if the presence of existential there allows the inference that the sentence has no topic, then the presence of a topic phrase is also in conflict with the structure of the rest of the sentence.

- (31) a. ?* There seems to be it on the table, a telephone.
 b. ?? A telephone, there seems to be Ø on the table.
 c. ?? There seems to be one on the table, a telephone.
 d. ? A telephone, there seems to be one on the table.

The oddness of the sentences of (31), in declining order, seems to come from the alleged topichood of the displaced and dislocated NP, and its definiteness of reference, or lack of it. If there occurs at all, then the NP a telephone has been displaced from possible subject/topic position, and hence cannot be perceived as topic. These properties found in the sentences of (31) show up more clearly by contrast with the sentences in (32), in which the dislocated NP is not the displaced subject. Since

the NP was not displaced from subject position, it might be a topic, and so may occur in topic phrases.

- (32) a. ? There seems to be a telephone on it, the table.
b. The table, there seems to be a telephone on it.
c. (?) The table, there seems to be a telephone on \emptyset .

While perhaps not splendidly felicitous, the sentences of (32) are certainly better than those of (31).

If the topic phrase and other features of syntactic structure and semantic content of the NP match in identifying a topic, there is no effect of redundancy, as would occur if two elements were combined which had identical or overlapping semantic content (cf. Grice, 1975, and Sadock, 1978). The topic-defining properties of marked sentence structures such as Raising to Object sentences and the extra material added in the topic phrase, are therefore not semantic in origin. These properties seem to be added on pragmatically, by inference about the speaker's intentions in choosing this structure when other equivalent but less complex structures are permitted by the grammar. These inferences cannot be cancelled because structural oppositions (as opposed to stylistic oppositions or rhetorical devices) cannot be cancelled--they continue to exist in the grammar of the language independently of the context of particular utterances. The result of combining indicators of topic which conflict is an infelicitous or pragmatically contradictory sentence. The sentences in (31), for example, are not necessarily deviant grammatically, but since they are contradictory in a pragmatic sense, it is hard to imagine of what use the sentence could be in any context.

Summary of Topic Defining Devices in English

In the preceding sections I have proposed ways of defining sentence topic that are particular to the grammar of English, and resemble other languages only to the extent that the grammar of English is similar to the grammar of another language. I have not said very much about universal factors, such as overall prior context, discourse topic, or emphasis indicated phonologically or with emphatic particles. These factors do contribute to the perception of topic, but context has been defined here only in the ways that it may reinforce or conflict with syntactic and semantic features. Ultimately, I am proposing a scale of psychologically salient features of the grammar of English such that Feature F is salient enough to define Constituent X as sentence topic in the absence of more salient elements. Some of the elements on the scale of grammatical properties are more salient than any contextual or emphatic feature, but those at the weaker end are not. I will assume some interaction between general factors such as these, and the specific, ranked grammatical features, but in this paper I am mainly concerned with isolating the latter.

In English, the grammatical definition of sentence topic is determined by the intersection of a NP semantic hierarchy of referring properties, given in (13), and a syntactic hierarchy of markedness, particularly given in (25). This scale (25) can be supplemented by two additional elements, one at the weakest end of the scale, the surface grammatical subject of a sentence in which no rule has changed the element in subject position, and another at the strongest end of the scale, which consists of NP phrases adjoined to a clause, as in examples (26) - (32). As the grammatical scale involves subjects primarily, it must be claimed that in English there is

something particularly salient about subjects, as opposed to the initial element of the sentence, which may subsume subjects not preceded by a preposed adverb.

By a generalized conversational inference or discourse strategy, as in (7), subjects are perceived as topics. The scale in (25) is based on subject positions and what they correspond to; the more marked the position, the greater the disparity between surface and underlying role. This characterization fits the NPs affected by Raising to Object or Exceptional Case-Marking in the framework of Chomsky (1981), as an underlying subject has anomalous accusative marking. Whether it actually remains a subject in spite of its case and object-like surface properties is a matter of controversy, but its markedness is undeniable. The most marked elements are outside the clause in which there is a coindexed element, and for this reason, the whole structure is marked. That is, if the speaker could choose to express some proposition using a single clause, but instead chooses an 'extra' phrase as well, then some inference about salience of the extra element is warranted.

Initial position in a sentence is not what is associated with syntactically defined sentence topics in English. It seems to be the case that initial position or proximity to it is characteristic of sentence topics in some languages (cf. the work of Firbas and others). But if this were true of English, then preposed adverbials should function as topics, and in that case, then NP types from the lower end of (13) should be possible in 'neutral' position, but not in preposed position. But the experimental evidence in Davison and Lutz (to appear) strongly suggests that preposed adverbials are not perceived as topics in the way that other, subject-like constituents are. And as (33) and (34) show,

It is possible to prepose 'poor topic' NPs without restriction of meaning or loss of acceptability:

- (33) a. The bomb is likely to go off at any moment.
- b. At any moment, the bomb is likely to go off.
- (34) a. They show their slides of the Alps at the slightest provocation/at the drop of a hat.
- b. At the slightest provocation/at the drop of a hat, they show their slides of the Alps.

The superlative in (34) and the generic adverbials in (33) and (34) are equally good in the a and b versions, with 'normal' stress and level pitch.

So the syntactic properties which may be used to define sentence topics in the absence of more salient features cluster around surface subject function, corresponding in some more marked cases to underlying objects, or to subject properties obscured by unusual case marking. This is so for all but the most marked structure, in which topic constituent is outside of clause structure. Of course, in this case it is hard to distinguish a dislocated NP from a preposed adverbial with high pitch and emphatic stress, but this may be the only instance of an initial adverbial topic (cf. Langacker, 1974):

If the function of sentence topic is associated with only one item, then there may be an explanation in the syntactic system of English for why adverbials are not perceived as topics, but subjects are. English lacks the rule of German which obligatorily inverts the subject with the first verbal element if some other constituent is preposed. For this reason, dislocated, topicalized or preposed constituents may precede a clause subject, in any number. If sentence topics were not uniquely identified by subject properties (out of other syntactic properties),

then no one of the preverbal constituents would be uniquely salient. But in languages like Norwegian which have an obligatory inversion rule, there can be only one preverbal constituent, which is perceived as the sentence topic (Faarlund, 1981). So, if the initial element in a main clause (or its pragmatic equivalent) is the subject, it is the topic; otherwise the subject and verb are inverted, and the preverbal constituent is a topic. It can be anaphoric to the context, or 'deictic,' serving as a connective to the context (Faarlund, 1981).

The definition of subject properties in surface structure needs to be made more explicit for English. Nominative case and verbal agreement are characteristic of subjects in finite clauses, and preverbal position in non-finite clauses. I would like to distinguish between inverted subjects in yes-no questions, which have topics corresponding to the topic of the equivalent assertion, from inverted subjects following there, which are not topics. So the correct stipulation seems to be that preverbal position is crucial for all subjects, unless there is no constituent preceding the verb.

There is another aspect of salience which determines what a sentence topic may be. So far I have defined topics in terms of clauses. A sentence consisting of many clauses could in principle have as many topics as it has overt subjects, by this definition; but generally this is not the case. A normal sentence, no matter how complex it may be syntactically, usually has at most one sentence topic, and if sentence topics are supposed to serve a discourse function of defining what the sentence is about, it would seem to follow that for such definition to be effective, there should be only one topic. So I will add to the scales of salience and

markedness the notion of clause salience, with main clauses as more salient than subordinate clauses.¹³ The major exception to this matching of syntactic subordination to pragmatic subordination is illustrated by the cases of Raising to Object and expressions of the type it seems that p, I think that p etc. Predicates of belief, perception and assertion, including the verbs that govern Raising, have a number of syntactic and pragmatic properties. They are in some sense redundant because their meaning is closely related to properties of assertions; constituents containing them may be perceived as pragmatically 'transparent,' so that I think p is pragmatically though not semantically equivalent to asserting p alone; constituents containing them may be parenthetical tags; and syntactic structures characteristic of main clauses may occur in their complements, under some circumstances (Hooper & Thompson, 1973, Green, 1976).

So for syntactically defined sentence topics, and possibly also for topics defined by referential properties, the topic of the whole sentence is the topic of the most pragmatically salient clause, generally the main clause. But note that an imperative main clause, lacking a subject, fails to define a topic, so that if there is a subordinate clause, its subject may be topic. So by this definition, (35)a lacks a topic, while (35)b - d have topics:

- (35) a. Shut the door.
- b. Consider this case to be equivalent to the preceding one.
- c. That door, shut it, please.
- d. That door, tell John he should shut it.

The raised subject in (35)b is a topic by both syntactic markedness and appropriate referring properties. That door in (35)c, and d is a topic

by the same properties, including very marked syntactic structure. Note that in this case, where there is a separate topic phrase, it does not matter whether it corresponds to a constituent of a main clause or a subordinate clause. If we want (35)a to have a topic, it would be defined by referring properties of the NP alone, not by syntax. If not, then a sentence such as (35)a might serve to introduce a topic into discourse, which then might occur as the topic of a succeeding sentence.

Conclusion

I have proposed in this paper that the notion of sentence topic is a pragmatic one, and does not correspond to a grammatical category defined by the grammar of English. But topics are defined by conversational implicature of a very regular and general kind on the basis of linguistic as well as contextual information. The information of a linguistic nature may include the syntactic properties of a sentence in which a potential topic NP occurs, as well as the semantic properties of a NP, particularly the ones which help to define its referent. The most regular and general inference is that the topic of a sentence is the subject of the superordinate clause, or the most pragmatically salient clause. The properties of grammatical subject which contribute to the definition of possible topics include preverbal position--so that post-verbal NPs in there constructions are not perceived as topics--verbal agreement morphology if no NP precedes the verb, and distinctive pronoun morphology. It is possible to cancel this inference in 'unmarked' syntactic structures, either by contextual information which defines some other constituent besides the subject as topic, or by the referring properties of the NP in subject position. If the referent of the subject NP is hard to determine from the information

given, then topic function is assigned to some other NP with more explicit referential properties, if there is one in the sentence. But in 'marked' syntactic contexts the subject-topic inference is harder to cancel. There are degrees of cancellability, since syntactic markedness is a matter of degree, determined by the extent to which the surface structure does not directly represent underlying logical relations. The more marked the syntactic structure is in which a possible topic NPs occurs, the more difficult it is to cancel the topic function of a NP assigned by inference. Where cancellation is ruled out by the markedness of the syntactic structure, then certain types of NP are generally ruled out in subject or other marked position. Their appearance in these positions is infelicitous, rather than ungrammatical, and empty of useful contribution to discourse.

This view of sentence topic in English, based on the syntax and lexicon of English, allows us to claim some psychological reality for the notion, without introducing the claim that the characterization of topics is operative in all circumstances and in all structures. By the definition and strategy I have offered here, the perception of topic is dependent on a number of factors, and is not integrated into the description of grammar. One might say that it is operative in the absence of other, stronger factors, including intonational patterns and emphatic stress, which I have not discussed here. That is, unless some cluster of factors defines a more salient element, the syntactic structure of a sentence and the content of noun phrases may define a sentence topic. This topic may agree with previous context, or merely be compatible with it; in more striking cases, it may conflict.

Topic-function can thus be separated from the rules of the grammar of the language (however they may be represented formally). The syntax of English can be regarded as autonomous, and without pragmatic conditions attached to syntactic rules or to semantic features. The choices of syntactic construction or NP type described in this paper are truly optional, in that I regard rules such as There Insertion and Passive as operating without regard to speaker beliefs or discourse contexts. These beliefs may be present, of course, and we may regard certain sentence structures as incompatible with discourse contexts or various beliefs. But to do this is not the same as incorporating pragmatic conditions into a syntactic/semantic system. The only sense in which this is done, in this paper, is in the sense that inferences are calculated on the basis of grammatical and semantic elements which form a system of oppositions in the language. Some import beyond literal meaning may legitimately be read into the choice of a more marked structure.

References

- Akmajian, A., & Kitagawa, C. Deep structure binding of pronouns and anaphoric bleeding. Language, 1976, 52, 61-77.
- Bever, T. A. The cognitive basis for linguistic structure. In J. R. Hayes (Ed.), Cognition and the development of language. New York: Wiley, 1970.
- Borkin, A. Raising to object position: A study of the syntax and semantics of clause merging. Unpublished dissertation, University of Michigan, 1974.
- Chafe, W. L. Language and consciousness. Language, 1974, 50, 111-133.
- Chafe, W. L. Givenness, contrastiveness, definiteness, subjects, topics and point of view. In C. N. Li (Ed.), Subject and topic. New York: Academic Press, 1976.
- Chomsky, N. Deep structure, surface structure and semantic interpretation. In D. Steinberg & L. Jakobovits (Eds.), Semantics. Cambridge: Cambridge University Press, 1971.
- Chomsky, N. Lectures on government and binding. Dordrecht: Foris Publications, 1981.
- Davison, A. Any as existential or universal? In J. van der Auwera (Ed.), The semantics of determiners. London: Croom Helm, 1980. (a)
- Davison, A. Peculiar passives. Language, 1980, 56, 42-66. (b)
- Davison, A., & Lutz, R. Measuring syntactic complexity relative to discourse (Technical Report). Urbana: University of Illinois, Center for the Study of Reading, to appear.

- Donellan, K. Reference and definite description. In D. Steinberg & L. Jakobovits, (Eds.), Semantics. Cambridge: Cambridge University Press, 1971.
- Faarlund, J. T. Obligatory fronting in a verb-initial language: An attempt at pragmatic syntax. Chicago Linguistics Society, 1981, 17, 45-58.
- Firbas, J. Non-thematic subjects in contemporary English. Travaux linguistiques de Prague, 1966, 2, 239-256.
- Green, G. M. Main clause phenomena in subordinate clauses. Language, 1976, 52, 382-397.
- Grice, H. P. Logic and conversation. In P. Cole & J. L. Morgan (Eds.), Syntax and semantics, Vol. 3. New York: Academic Press, 1975.
- Gundel, J. K. Role of topic and comment in linguistic theory. Ohio State University dissertation. Bloomington: Indiana University Linguistics Club, 1977.
- Halliday, M. A. K. Notes on transitivity and theme in English: II. Journal of Linguistics, 1967, 3, 199-244.
- Haviland, S. E., & Clark, H. H. What's new? Acquiring new information as a process in comprehension. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 512-521.
- Hooper, J. B., & Thompson, S. A. On the applicability of root transformation. Linguistic Inquiry, 1973, 4, 465-498.
- Horn, L. R. Richard redux: New evidence for a role of subject copying. Paper presented at the 1980 Winter meeting of the Linguistic Society of America, 1980.
- Kantor, R. N. The management and comprehension of discourse connection by pronouns. Ohio State University dissertation, 1977.

- Keenan, E. L. Towards a universal definition of 'subject.' In C. N. Li (Ed.), Subject and topic. New York: Academic Press, 1976.
- Langacker, R. Movement rules in functional perspective. Language, 1974, 50, 630-674.
- Li, C. N. (Ed.) Subject and topic. New York: Academic Press, 1976.
- Li, C. N., & Thompson, S. A. Subject and topic: A new typology of language. In C. N. Li (Ed.), Subject and topic. New York: Academic Press, 1976.
- Milsark, G. Towards an explanation of certain peculiarities of the existential construction in English. Linguistic Analysis, 1977, 3, 1-30.
- Olson, D., & Filby, N. On the comprehension of active and passive sentences. Cognitive Psychology, 1972, 3, 361-381.
- Prince, E. Towards a taxonomy of given-new information. In P. Cole (Ed.), Radical pragmatics. New York: Academic Press, 1980.
- Rando, E. N., & Napoli, D. J. Definites in there sentences. Language, 1978, 54, 300-313.
- Reinhart, T. Pragmatics and linguistics: An analysis of sentence topics. Philosophica, 1981, 27, 53-94.
- Sadock, J. M. On testing for conversational implicature. In P. Cole (Ed.), Syntax and semantics, Vol. 9. New York: Academic Press, 1979.
- Schmerling, S. F. A reexamination of 'normal stress.' Language, 1974, 50, 66-73.
- Sinha, A. K. How passive are passives? Chicago Linguistic Society, 1974, 10, 631-642.
- Strawson, P. F. Identifying reference and truth values. In D. Steinberg & L. Jakobovits (Eds.), Semantics. Cambridge: Cambridge University Press, 1971.

- Vendler, Z. Linguistics and philosophy. Ithaca: Cornell University Press, 1967.
- Wanner, E., & Maratsos, M. An ATN approach to comprehension. In M. Halle, J. Bresnan, & G. A. Miller (Eds.), Linguistic theory and psychological reality. Cambridge: M.I.T. Press, 1978.
- Ziff, P. The non-synonymy of actives and passives. Philosophical Review, 1966, 75, 226-232.

Footnotes

The research reported herein was supported in part by the National Institute of Education under Contract No. HEW-NIE-C-400-76-0116. Earlier versions of this paper were presented as a paper at the Tenth NWave meeting, University of Pennsylvania, October 1981, and as colloquia at the University of Arizona and Michigan State University. I am grateful to people who were present on these occasions for discussion and comments which have been particularly helpful to me at various stages of writing and revising this paper. I want to thank especially Barbara Abbott, Adrian Akmajian, Michael Harnish, Adrienne Lehrer, Jerry Morgan and Ellen Prince (who is also hereby thanked for extraordinary aid on the morning of October 24, 1981).

¹On this point, I think that the point made by Reinhart (1981) is compelling: within a sentence, any second mention of an individual automatically makes the referring expression a topic phrase, if topics represent old information. But second mention in a sentence usually does not have this effect. The alternative characterization of sentence topics, what the sentence is 'about,' is a more realistic description of the function of sentence topics, though not a very explicit criterion for picking out the sentence topic in any particular case. But the function of defining 'aboutness' explains better than the property of being 'old information' why sentences do not have multiple sentence topics (I do not exclude the possibility that they may have multiple discourse topics, following Reinhart). If the property of being 'old information' is defined by prior mention, or context, then it could be true of many constituents in a sentence that

they represent old information. Thus topics defined in this way could match many constituents in a sentence. But as Adrienne Lehrer has pointed out (p.d.), communication is more efficient and easily related to discourse if what the sentence is about is just one entity, designated by one sentence constituent. Of course, a sentence may be about different things on different occasions of utterance; topic-function is not uniquely and constantly assigned to grammatical units in the same way that grammatical roles are.

²There is clearly something pragmatically misleading about specifying information about something which either does not exist, or is not identifiable by virtue of the linguistic expression in subject/topic position, and this conflict of semantic, syntactic and pragmatic information would be hard to assign a specific truth value; predicated information may be evaluated relative to what is known about a certain entity. The characterization of subjects as topics is not explicitly made by Strawson in the article cited, but it is implicit in his use of 'center of interest,' and active and passive sentences as examples.

³In the experiment reported in Davison and Lutz (to appear), reaction time was measured for reading the second of a sequence of two sentences. In the second sentence, there were two possibilities: either an 'optional' transformation had applied, or it did not, though the sentence was a counterpart of the first version. The transformations in question were Passive, There Insertion, Adverb Preposing, Raising to Subject and Raising to Object. The 'target' sentence was preceded by a context sentence, which could be one of three kinds: (i) neutral; (ii) biased to the transformed version and against the untransformed version; or (iii) biased to the

untransformed version and, where possible, against the transformed version. Bias was created by making an element of the context match what was supposed to be in topic position, by class membership or alternative description. In the Adverb Preposing cases, there was no difference in reaction time for the preposed and unpreposed versions (ii) and (iii) in the second kind of context (i), in which the context is biased towards the subject and not the preposed adverbial.

- (i) (Context) Shopping in other countries can be frustrating for Americans.
- (ii) (Untransformed target) Most English shops are not open after 6 p.m.
- (iii) (transformed target) After 6 p.m., most English shops are not open.

These results strongly suggest that initial elements-which would include preposed adverbials-are not topics necessarily. Reaction times should have been much longer for the transformed version if preposed adverbials could be topics when the context is biased towards the subject. In fact the reverse is found; in the context biased towards the preposed adverbial, actual reaction times were somewhat longer for the preposed version, as though the linking of context with preposed adverbial, rather than subject, was somewhat misleading.

⁴The active and passive versions of the 'same' sentence are not always equally preferred in a given discourse context. Sinha (1974) provides a number of examples of passive sentences used in contexts where the corresponding active version would be strange or certainly inferior to the passive version. He claims that such cases indicate that there is an

Inherent semantic difference between the active and passive constructions. But I argue in this paper that they are equivalent in meaning but are perceived differently in some contexts because the topic-subject inference applies to different constituents. As Ziff (1966) observed, active and passive sentences are about different things.

⁵In these examples, I regard sit as an intransitive verb which never takes a true direct object. It contrasts with the causative and transitive counterpart seat, like fall, fell. For arguments that passive sentences with sit Prep are derived from sources with intransitive verbs, see Davison (1980-b).

⁶This case contrasts with the failure-of-reference cases discussed by Strawson (1971), and explained in terms of contradiction between the semantic content of the NP and its lack of referent in the world. Names like Uncle Forrester and definite descriptions such as the man who broke the bank at Monte Carlo imply at least pragmatically that there exists a referent, (and the referent is known to the speaker and possibly also to the hearer. There and other NPs which are semantically empty, or nearly so, and which are not used as names, do not give any information about a purported referent. They may be used in subject position with cancellation of the topic inference, rather than with contradiction.

⁷Sentence (10)b and (11)b are marked as strange in the existential reading with a non-specific indefinite reading for the NP. These sentences are much better in a discourse context introducing referents for the NPs, making them specific: Where are the three flies? or There are three waiters and I know the whereabouts of two of them. These contexts and readings are independent of the point being made here.

⁸The pattern of reaction times reported in Davison and Lutz (to appear) for Raising to Object sentences indicate that the lower clause subject functions as the topic of the sentence overall. Reaction time is lower for both versions of the sentence if the context sentence is linked with the lower clause subject, and are high otherwise.

⁹As the preceding note suggests, the lower clause NP may be perceived as sentence topic even if Raising does not apply, because of the pragmatic transparency of the higher clause. (The class of verbs believe, report, consider is semantically related to various components of assertions, so that the superordinate clause containing one of these verbs may be pragmatically, though not semantically, disregarded. I am grateful to David Dowty for discussion clarifying this point.)

¹⁰Another example of the persistence of what is conveyed by marked structures may be seen by comparing subjectless imperative sentences with other sentences having subjects which may be used in an imperative sense.

- (i) You will open the transom, but I am not asking you to open the transom.

There is no contradiction in (i), in spite of the fact that (ii) may be used like (iii):

- (ii) You will open the transom(!)

- (iii) Open the transom.

- (iv) #Open the transom, but I'm not asking you to open the transom.

If (iii) is meant to be taken as an imperative, and not as a sentence fragment with a specific prior discourse context, then (iv) is contradictory. The difference is that the tenseless subjectless structure is 'marked' while the subject-tensed verb combination in (ii) is not. The inference

that it is meant as a directive is cancellable, while the directive quality of (iv) is not.

¹¹ Horn (1980) argues that the higher clause position in (i) conventionally implies that the NP occupying the subject position has a referent and the referent is existent.

(i) Susan looks like she's going to say something.
The output of the rule of 'Richard,' or Subject Copying, illustrated here is 'marked' for the same reason as the output of Raising to Subject, with the additional factor that there is a pronoun copy instead of a null subject in the lower clause. The structure in (i), therefore is not part of the class to which Equi-NP Deletion and Raising to Subject structures belong. As Horn shows, the rule may not place an indefinite with non-specific reference in subject position, and other NP types are also restricted, as for Raising to Object:

(ii) The smallest particle looks like it's going to be discovered.

(iii) The jig looks like it's up.

(iv) Any doctor sounds like he can give you a prescription for the medicine.

(v) It sounds like any doctor can give you a prescription for the medicine.

(ii) - (iv) are strange, but note that (v) is not odd, and is a way that the content of (iv) can be expressed. The oddity of (iv) is therefore due to syntactic structure. What is called syntactic markedness here corresponds to Horn's conventional implicature associated with subject position.

¹²As Akmajian and Kitagawa (1976) point out, NPs containing quantifiers undergo relativization (and Raising to Object AD) but make poor topics in dislocated or topicalized phrases. Nevertheless, some quantified expressions, with definite articles particularly, can be topicalized.

- (i)

All those books, Two of those books, Most of those books, A few of those books,
--

 I can't stand (them).

Quantified expressions like these should be able to function as topics, since they give some information about the intended referents, relative to the set of possible referents--i.e., none, some, a certain quantity, etc.

¹³As for conjoined sentence structures, I know of no reason for assuming that the syntactic construction defines one element as more pragmatically salient than the other. So the sentence topics of each will be perceived as if the two sentences were adjacent in discourse.